PLAY-THE-WINNER RULE AND INVERSE SAMPLING FOR SELECTING THE BEST OF $k \geq 3$ BINOMIAL POPULATIONS.

by

Milton Sobel* and George H. Weiss+

University of Minnesota Technical Report 126

Imperial College, London

July, 1969

*This author was supported in part by a National Institutes of Health Special Fellowship and in part by National Science Foundation Grant GP-9018 while on leave from the University of Minnesota, Minneapolis, Minnesota.

+On leave from National Institutes of Health, Bethesda, Maryland.

1. Introduction. The effect of the play-the-winner (PW) sampling rule when used with different termination rules for selecting the best of k independent binomial populations has already been considered for $k = 2$ in [5] and [6]. The best population is defined to be the one with the highest probability p of success on a single trial. In this paper we consider $k \geq 3$ and consider a cyclic variation of the play-the-winner (PWC) sampling rule when the termination rule is "inverse sampling." In other words for $k = 3$ we sample (say) population A until we observe a failure, then switch to (say) B and keep on with B until its first failure, and then switch to C and observe C until its first failure; the cycle then gets repeated and when one of the players gets r successes we stop and select it as best. At the outset we use randomization, i.e., we order (for sampling) the 3 populations giving probability 1/6 to each possible ordering.

As in [5] and [6], we specify constants P^* and Δ^* with $\frac{1}{k} < P^* < 1$ and $0 < \Delta^* < 1$ and require a procedure R for which the probability of a correct selection (CS) satisfies

$$(1.1) \qquad P\{CS\} \geq P^* \qquad \text{whenever} \quad \Delta \geq \Delta^*;$$

here Δ denotes the true difference between the largest and second largest p-value.

The PW sampling rule was first suggested by Robbins and studied in connection with the 2-arm bandit problem by several authors (see the references in [5]). We shall also be interested in the vector-at-a-time (VT) sampling rule which was extensively employed in sequential ranking problems in [1]. Here we take k observations at each stage, one from each population, and refer to these k-tuples as vectors. Then the total number of trials (or observations) at termination must be k times the number of vectors observed. We denote the inverse sampling procedure that uses PWC sampling by R_I and the one that uses VT sampling by R_I' as in [6].

It is shown that for large values of r the procedure R_I with the PW sampling rule is uniformly preferable to the procedure R_I' with the VT sampling rule in the sense that (with r chosen to satisfy (1.1) in both) the former requires a smaller expected total number of observations $E\{N\}$ for all parameter points with $\Delta > 0$. An expected loss (or risk) function is defined for any procedure R by

$$E\{L|R\} = \sum_{i=1}^{k} (p_1 - p_i)\, E\{N_i|R\}, \tag{1.2}$$

where p_1 is the largest of the p_i and N_i is the number of observations from the population with success parameter p_i. Using (1.2) to compare procedures rather than $E\{N\}$, the procedure R_I is again uniformly preferable to procedure R_I' for large values of r. The value of r above which these results hold is estimated, but no bound on the accuracy of this estimate is given.

2. The Procedure R_I: Exact Results.

Under inverse sampling we stop as soon as one population has r successes and select it as best; the integer $r > 0$ is predetermined so that (1.1) is satisfied. We wish to find the probability of a correct selection $P\{CS|R_I\}$ under the procedure R_I, which uses PW sampling

Let A_1 denote the best population, A_2 the one following A_1 in the initial randomization, etc. (continuing in cyclic order) and let $S(A_i) = S_i$ denote the current number of successes for A_i, so that $r - S_i = T_i$ is the number of successes A_i needs to be selected as best. Let $\underset{\sim}{T} = (T_1, T_2, \ldots, T_k)$. We define probabilities $U_i(\underset{\sim}{m}) = U_i(m_1, m_2, \ldots, m_k)$ for $i = 1, 2, \ldots, k$ by

$$(2.1) \qquad U_i(\underset{\sim}{m}) = P\{CS \mid \underset{\sim}{T} = \underset{\sim}{m} \text{ and the next observation is on } A_i\}$$

and use p_i to denote the single-trial probability of success for population A_i $(i = 1, 2, \ldots, k)$. From the PWC sampling rule we obtain the k recursions $(i = 1, 2, \ldots, k)$

$$(2.2) \qquad U_i(\underset{\sim}{m}) = p_i U_i(m_1, m_2, \ldots, m_i - 1, \ldots, m_k) + q_i U_{i+1}(\underset{\sim}{m})$$

where $U_{k+1} \equiv U_1$ and boundary conditions are given by

$$(2.3) \qquad \begin{aligned} &U_1(0, m_2, \ldots, m_k) &&= 1 \quad \text{if } m_j > 0 \text{ for } j \neq 1 \\ &U_i(m_1, \ldots, m_{i-1}, 0, m_{i+1}, \ldots, m_k) &&= 0 \quad \text{if } m_j > 0 \text{ for } j \neq i. \end{aligned}$$

To find a solution of (2.2) satisfying (2.3) we use generating functions $V_i = V_i(\underset{\sim}{x}) = V_i(x_1, x_2, \ldots, x_k)$ defined by

$$(2.4) \qquad V_i = \sum_{m_1=1}^{\infty} \cdots \sum_{m_k=1}^{\infty} U_i(\underset{\sim}{m}) x_1^{m_1} x_2^{m_2} \cdots x_k^{m_k} \qquad (i = 1, 2, \ldots, k).$$

It is readily verified that (2.2) leads to

$$(2.5) \qquad \begin{aligned} &(1-p_1x_1)V_1 - q_1V_2 = p_1x_1 \prod_{j=2}^{k} \left(\frac{x_j}{1-x_j}\right) \\ &(1-p_ix_i)V_i - q_iV_{i+1} = 0 \qquad \text{for } i = 2, \ldots, k \end{aligned}$$

(where $V_{k+1} \equiv V_1$) and hence, letting $D = (1-p_1x_1)(1-p_2x_2) \cdots (1-p_kx_k) - q_1q_2 \cdots q_k$, we obtain

$$(2.6) \qquad \begin{aligned} V_1 &= \frac{p_1x_1}{D} \prod_{j=2}^{k} \left(\frac{x_j(1-p_jx_j)}{1-x_j}\right) \\ V_i &= \frac{p_1x_1}{D} \left[\prod_{j=2}^{k} \left(\frac{x_jq_j}{1-x_j}\right)\right] \prod_{\alpha=2}^{i-1} \left(\frac{1-p_\alpha x_\alpha}{q_\alpha}\right) \qquad (i = 2, 3, \ldots, k); \end{aligned}$$

in this paper products with no factors are taken equal to one. Since we use randomization with equal probabilities $\frac{1}{k}$ for each population at the outset, it follows that

$$(2.7)\qquad P\{CS|R_I\} = \frac{1}{k}\sum_{i=1}^{k} U_i(\underset{\sim}{r}) \qquad ,$$

i.e., the coefficient of $x_1^r x_2^r \cdots x_k^r$ in $\frac{1}{k}\sum_{i=1}^{k} V_i$, where $\underset{\sim}{r} = (r, r, \ldots, r)$ and r is chosen to satisfy (1.1).

To get an explicit expression for (2.7) we use the expansion of $1/D$ given by

$$(2.8)\qquad \frac{1}{D} = \sum_{i=0}^{\infty} \frac{(q_1q_2\cdots q_k)^i}{[\prod_{j=1}^{k}(1-p_jx_j)]^{i+1}} = \sum_{\alpha=0}^{\infty}\sum_{\beta=0}^{\infty}\cdots\sum_{\omega=0}^{\infty}(p_1x_1)^{\alpha}(p_2x_2)^{\beta}\cdots(p_kx_k)^{\omega} \;.$$

$$\cdot \sum_{i=0}^{\infty}\binom{i+\alpha}{\alpha}\binom{i+\beta}{\beta}\cdots\binom{i+\omega}{\omega}(q_1q_2\cdots q_k)^i \;.$$

Using the well-known identity for the incomplete beta function (see e.g. [5])

$$(2.9)\qquad q^r\sum_{j=0}^{s-1}\frac{\Gamma(r+j)}{\Gamma(r)j!}p^j = I_q(r, s) = p^s\sum_{j=r}^{\infty}\frac{\Gamma(s+j)}{\Gamma(s)j!}q^j$$

(where the first equality holds for any real $r > 0$ and the second for any real $s > 0$), we find from (2.6) and (2.8) that the coefficient of $(x_1x_2\cdots x_k)^r$ in V_1 is

$$(2.10)\qquad U_1(\underset{\sim}{r}) = p_1^r\sum_{i=0}^{\infty}\binom{i+r-1}{i}q_1^i\prod_{j=2}^{k}I_{q_j}(i, r) \qquad ,$$

where $I_q(0, r) = 1 = 1-I_p(r, 0)$ for $r > 0$. From V_α with $\alpha \geq 2$ we obtain

$$(2.11)\qquad U_\alpha(\underset{\sim}{r}) = p_1^r\sum_{i=0}^{\infty}\binom{i+r-1}{i}q_1^i\left[\prod_{j=2}^{\alpha-1}I_{q_j}(i, r)\right]\left[\prod_{j=\alpha}^{k}I_{q_j}(i+1, r)\right] .$$

Hence by (2.7) we can use (2.10) and (2.11) to write

$$(2.12) \qquad P\{CS|R_I\} = \frac{1}{k} E_r\left\{\left[\prod_{j=2}^{k} I_{q_j}(X, r)\right] + \sum_{\alpha=2}^{k}\left[\prod_{j=2}^{\alpha-1} I_{q_j}(X, r)\right]\left[\prod_{j=\alpha}^{k} I_{q_j}(X+1, r)\right]\right\},$$

where the random variable X has the negative binomial distribution with index $r > 0$, success parameter p_1 and mean rq_1/p_1 (cf. 2.9 above).

Similar calculations are used to find the expected number of observations $E\{N_i|R_I\}$ on A_i under procedure R_I and the sum of these is the expected total number of observations. For fixed i, let

$$(2.13) \qquad S_j(\underset{\sim}{m}) = E\{N_i|\underset{\sim}{T} = \underset{\sim}{m} \text{ and the next trial is on } A_j\} \quad (j = 1, 2, \ldots, k).$$

As in (2.2) we obtain the recursions

$$(2.14) \qquad S_j(\underset{\sim}{m}) = p_j S_j(m_1, m_2, \ldots, m_i-1, \ldots, m_k) + q_j S_{j+1}(\underset{\sim}{m}) + \delta_{ji} \quad ,$$

where $\delta_{ji} = 1$ for $j = i$ and zero otherwise, and $S_{k+1} \equiv S_1$. The boundary conditions are

$$(2.15) \qquad S_j(m_1, \ldots, m_{j-1}, 0, m_{j+1}, \ldots, m_k) = 0 \qquad \text{if } m_\alpha > 0 \text{ for } \alpha \neq j.$$

The desired result is obtained by finding

$$(2.16) \qquad E\{N_i|R_I\} = \frac{1}{k}\sum_{j=1}^{k} S_j(\underset{\sim}{r}).$$

Using the generating functions as in (2.5), we define $T_i = T_i(\underset{\sim}{x})$ and obtain

$$(2.17) \qquad \begin{aligned} T_i &= \frac{1}{D}\prod_{\alpha=1}^{k}\left(\frac{x_\alpha}{1-x_\alpha}\right)\prod_{j\neq i}(1-x_j p_j) \\ T_j &= T_i \prod_{\alpha=j}^{i-1}\left(\frac{q_\alpha}{1-x_\alpha p_\alpha}\right) && \text{for } j < i \\ T_j &= T_i \prod_{\alpha=j}^{k}\left(\frac{q_\alpha}{1-x_\alpha p_\alpha}\right)\prod_{\beta=1}^{i-1}\left(\frac{q_\beta}{1-x_\beta p_\beta}\right) && \text{for } j > i \quad , \end{aligned}$$

where D, defined as above, is expanded in (2.8). From (2.17) we obtain for $j \le i$ and $j > i$, respectively,

$$
(2.18)\qquad
\begin{aligned}
S_j(\underset{\sim}{r}) &= \frac{1}{q_i} \sum_{\alpha=0}^{\infty} I_{q_i}(\alpha+1, r) \left[\prod_{\beta=j}^{i-1} I_{q_\beta}(\alpha+1, r)\right]\left[\prod_{\substack{\gamma<j \\ \text{and} \\ \gamma>i}} I_{q_\gamma}(\alpha, r)\right] \\
S_j(\underset{\sim}{r}) &= \frac{1}{q_i} \sum_{\alpha=0}^{\infty} I_{q_i}(\alpha+1, r) \left[\prod_{\substack{\beta<i \\ \text{and} \\ \beta\ge j}} I_{q_\beta}(\alpha+1, r)\right]\left[\prod_{\gamma=i+1}^{j-1} I_{q_\gamma}(\alpha, r)\right]
\end{aligned}
$$

By (2.16) the average of these k quantities ($j = 1, 2, \ldots, k$) in (2.18) is $E\{N_i | R_I\}$ and the sum ($i = 1, 2, \ldots, k$) of these k averages is the exact expected value $E\{N | R_I\}$ of the total number of observations required by procedure R_I.

3. Approximations and the Determination of r for Procedure R_I.

Since the incomplete beta function is decreasing in the first argument and increasing in the second argument, we can get bounds on the $P\{CS | R_I\}$ in (2.12) with the same asymptotic value for $r \to \infty$ by replacing X by X +1 or vice versa, obtaining

$$
(3.1)\qquad E_r\left\{\prod_{j=2}^{k} I_{q_j}(X+1, r)\right\} < P\{CS | R_I\} < E_r\left\{\prod_{j=2}^{k} I_{q_j}(X, r)\right\}.
$$

Let X_{p_i} ($i = 1, 2, \ldots, k$) denote independent negative binomial chance variables with success parameter p_i and common index r, so that in (3.1) $X = X_{p_1}$. Then, letting $\Delta_{1j} = p_1 - p_j$, we have for large r

$$(3.2)\qquad P\{CS|R_I\} \sim P\{X_{p_1} < X_{p_j}\ (j = 2, 3, \ldots, k)\}$$

$$= P\left\{\frac{X_{p_j} - \frac{rq_j}{p_j}}{\sqrt{rq_j}/p_j} > \left(\frac{X_{p_1} - \frac{rq_1}{p_1}}{\sqrt{rq_1}/p_1}\right)\frac{p_j}{p_1}\sqrt{\frac{q_1}{q_j}} - \frac{\Delta_{1j}\sqrt{r}}{p_1\sqrt{q_j}} \quad (j = 2, 3, \ldots, k)\right\}$$

$$\sim \int_{-\infty}^{\infty} \prod_{j=2}^{k}\left[1-\Phi\left(\frac{xp_j\sqrt{q_1} - \Delta_{1j}\sqrt{r}}{p_1\sqrt{q_j}}\right)\right] d\Phi(x)$$

$$= \int_{-\infty}^{\infty} \prod_{j=2}^{k} \Phi\left(\frac{xp_j\sqrt{q_1} + \Delta_{1j}\sqrt{r}}{p_1\sqrt{q_j}}\right) d\Phi(x),$$

where $\Phi(x)$ is the standard normal distribution function.

For the first step in the minimization of (3.2) subject to the conditions $\Delta_{1j} \geq \Delta^*$ $(j = 2, 3, \ldots, k)$, we note from (2.12) that the exact PCS is strictly decreasing in each p_j for $j \geq 2$ and hence we set $p_j = p_2$ and $\Delta_{1j} = \Delta^*$ for $j \geq 2$. Then we can write (3.2) for $\rho > 0$ in the form

$$(3.3)\qquad \text{Min } P\{CS|R_I\} \sim \int_{-\infty}^{\infty} \Phi^{k-1}\left(\frac{x\sqrt{\rho} + H}{\sqrt{1-\rho}}\right) d\Phi(x) = A_{k-1}(\rho, H)$$

where the last equality defines $A_{k-1}(\rho, H)$,

$$(3.4)\qquad H = \frac{\Delta^*\sqrt{r}}{\sqrt{q_1p_2^2 + q_2p_1^2}} \quad \text{and} \quad \rho = \frac{q_1p_2^2}{q_1p_2^2 + q_2p_1^2} \quad .$$

Since H is the same as for $k = 2$ in [6], (3.3) also holds for $k = 2$.

For the second part of the minimization we do not get an exact result for $k \geq 3$, only an approximate result. For this purpose we write the $P\{CS|R_I\}$ in (3.2) in the form

$$(3.5) \qquad P\{CS|R_I\} = P\left\{\frac{X_{p_1} - X_{p_j} - r\left(\frac{q_1}{p_1} - \frac{q_2}{p_2}\right)}{\sqrt{r}\sqrt{q_2 p_2^{-2} + q_1 p_1^{-2}}} < H \quad (j = 2, 3, \ldots, k)\right\}$$

where H and for $k \geq 3$ the correlation ρ between any two differences, such as $X_{p_1} - X_{p_2}$ and $X_{p_1} - X_{p_3}$, are both given by (3.4). We obtain an approximate minimum of (3.5) by minimizing H in (3.5) and disregarding (at first) the fact that ρ is also varying. This gives the same result as for $k = 2$ in equation (4.5) of [6], namely

$$(3.6) \qquad p_1 = \frac{2}{3} + \frac{\Delta^*}{2} \; ; \quad p_2 = p_3 = \cdots = p_k = \frac{2}{3} - \frac{\Delta^*}{2} \quad ,$$

obtained by maximizing the square of the denominator of H in (3.4); the details are given in [6]. Putting (3.6) in the expression for ρ in (3.4) gives

$$(3.7) \qquad \rho = \frac{1}{2} - \frac{3\Delta^*}{2} + \mathcal{O}\{(\Delta^*)^2\} \quad ;$$

this indicates that for $\rho = \frac{1}{2}$ we get a first approximation and that a correction term might be desirable.

This type of situation where a result is exact for $k = 2$ and is only approximately true for $k \geq 3$, the closeness of the approximation depending on how small Δ^* is, was also obtained in selecting the best of k binomial populations with VT sampling in [4].

For the correction term we expand the right side of (3.3) about $\rho = \frac{1}{2}$ and use (3.6) and (3.7). For this purpose we let $\varphi(x)$ denote the standard normal density and use the

<u>Lemma</u>: For $k \geq 3$, H fixed and positive $\rho < 1$

$$(3.8) \qquad \frac{d}{d\rho} A_{k-1}(\rho, H) = \frac{(k-1)(k-2)\,\varphi(H)\,\varphi\left(H\sqrt{\frac{1-\rho}{1+\rho}}\right)}{2\sqrt{1+\rho}} A_{k-3}\left(\frac{\rho}{1+2\rho}, H\sqrt{\frac{1-\rho}{(1+\rho)(1+2\rho)}}\right).$$

<u>Proof</u>: Differentiation under the integral sign gives for the left member M of (3.8)

$$(3.9) \qquad M = \frac{(k-1)}{2(1-\rho)\sqrt{\rho(1-\rho)}} \int_{-\infty}^{\infty} (x+H\sqrt{\rho})\, \varphi(x)\, \varphi\left(\frac{x\sqrt{\rho}+H}{\sqrt{1-\rho}}\right) \Phi^{k-2}\left(\frac{x\sqrt{\rho}+H}{\sqrt{1-\rho}}\right) dx .$$

If we 'complete the square' and make the appropriate transformation $y = (x+H\sqrt{\rho})/\sqrt{1-\rho}$, then we obtain

$$(3.10) \qquad M = \frac{(k-1)\,\varphi(H)}{2\sqrt{\rho}} \int_{-\infty}^{\infty} y\, \Phi^{k-2}(y\sqrt{\rho} + H\sqrt{1-\rho})\, d\Phi(y).$$

Integrating-by-parts we obtain

$$(3.11) \qquad M = \frac{(k-1)(k-2)}{2} \int_{-\infty}^{\infty} \varphi(y)\, \varphi(y\sqrt{\rho} + H\sqrt{1-\rho})\, \Phi^{k-3}(y\sqrt{\rho} + H\sqrt{1-\rho})\, dy.$$

If we 'complete the square' in (3.11) and make the appropriate transformation $y + H\sqrt{\rho(1-\rho)}/(1+\rho) = Z/\sqrt{1+\rho}$, then we obtain the desired result on the right side of (3.8).

If we set $\rho = \frac{1}{2}$ in the right side of (3.8) we obtain

$$(3.12) \qquad \frac{(k-1)(k-2)}{\sqrt{6}}\, \varphi(H)\, \varphi(H/\sqrt{3}) \int_{-\infty}^{\infty} \Phi^{k-3}\left(\frac{\frac{x}{2} + \frac{H}{\sqrt{6}}}{\sqrt{3}/2}\right) d\Phi(x) \quad ,$$

where $H = \Delta^{*}\sqrt{27r/8}$ and the integral in (3.12) is in the same form as in (3.3) with new values (H_1, ρ_1) given by

$$(3.13) \qquad H_1 = \frac{H}{\sqrt{6}} = \Delta^{*}\sqrt{\frac{27r}{48}} \quad ; \qquad \rho_1 = \frac{\rho}{2} = \frac{1}{4} \; .$$

Using (3.7) our minimum $P\{CS|R_I\}$ can now be written for small Δ^{*} as

$$(3.14) \qquad \text{Min } P\{CS|R_I\} \sim A_{k-1}(\tfrac{1}{2}, H) + (\rho-\tfrac{1}{2})\, \frac{(k-1)(k-2)}{\sqrt{6}}\, \varphi(H)\, \varphi\left(\frac{H}{\sqrt{3}}\right) A_{k-3}\left(\frac{1}{4}, \frac{H}{\sqrt{6}}\right)$$

$$\sim A_{k-1}(\tfrac{1}{2}, H) - \frac{3\Delta^{*}}{2}\, \frac{(k-1)(k-2)}{\sqrt{6}}\, \varphi(H)\, \varphi\left(\frac{H}{\sqrt{3}}\right) A_{k-3}\left(\frac{1}{4}, \frac{H}{\sqrt{6}}\right)$$

where $H = \Delta^* \sqrt{27r/8}$. We now get a first approximation for r by setting

$$A_{k-1}(\tfrac{1}{2}, H) = P^* \tag{3.15}$$

which can be done with existing tables (e.g. [2] or [3]), and then use the last expression in (3.14) to make minor corrections in r.

If $\lambda = \lambda(P^*)$ is the value of H that satisfies (3.15) then for small Δ^* we have as in equation (4.7) of [6]

$$r \sim \frac{8}{27}\left(\frac{\lambda}{\Delta^*}\right)^2 \tag{3.16}$$

Thus the results for $k \geq 3$ are of the same form as for $k = 2$.

For example, if we take $k = 3$, $P^* = .90$ and $\Delta^* = .10$ we find, using [2] or [3] that the solution of (3.15) is $H = 1.58$. Setting $H = 1.60$ in (3.14) we obtain $.9043 - .0035 = .9008$ and for $H = 1.59$ we obtain $.9023 - .0036 = .8987$, so that $H = 1.60$ is the closer value. Then, using (3.16) with $\lambda = 1.60$, we estimate r to be 75.85, so that $r = 76$ is required to satisfy (1.1) for the given values of k, P^* and Δ^*. If we had used the uncorrected $H = 1.58$ in (3.16), we would have obtained $r = 74$.

For the expected number of trials we use (2.18) and the fact that for $r \to \infty$

$$I_q(\alpha, r) = I_q(\alpha+1, r) + \binom{j+r-1}{j} p^r q^j = I_q(\alpha+1, r)[1 + o(1)]. \tag{3.17}$$

Thus we can approximate $S_j(\underset{\sim}{r})$ in (2.18) for all j and $E\{N_i|R_I\}$ by

$$E\{N_i|R_I\} \sim \frac{1}{q_i} \sum_{\alpha=0}^{\infty} \left[\prod_{\beta=1}^{k} I_{q_\beta}(\alpha+1, r)\right] \sim S_j(\underset{\sim}{r}) \qquad (j = 1, 2, \ldots, k), \tag{3.18}$$

where the infinite sum does not depend on i or j. Hence for $r \to \infty$ (or $\Delta^* \to 0$) we also have

$$E\{N|R_I\} \sim \left(\sum_{i=1}^{k} \frac{1}{q_i}\right) \sum_{\alpha=0}^{\infty} \left[\prod_{\beta=1}^{k} I_{q_\beta}(\alpha + 1, r)\right] . \tag{3.19}$$

Using the second identity in (2.9), the infinite sum Z in (3.19) becomes

$$(3.20) \qquad Z = p_1^r \sum_{j=0}^{\infty} \binom{j+r-1}{j} q_1^j \sum_{\alpha=0}^{j-1} [1-I_{p_2}(r, \alpha+1)][1-I_{p_3}(r, \alpha+1)] \; .$$

Multiplying out the bracketed expressions, we use lemma 1 of [2] for the two middle terms and obtain

$$(3.21) \qquad Z = \frac{rq_1}{p_1} + \frac{r}{p_2} E_r\{I_{p_2}(r+1, X)\} - \frac{r}{p_1} E_{r+1}\{I_{p_2}(r, X)\}$$

$$+ \frac{r}{p_3} E_r\{I_{p_3}(r+1, X)\} - \frac{r}{p_1} E_{r+1}\{I_{p_3}(r, X)\}$$

$$+ E_r\{\sum_{\alpha=0}^{X-1} I_{p_2}(r, \alpha+1) I_{p_3}(r, \alpha+1)\}.$$

For $p_1 > p_j$ $(j = 2, 3, \ldots, k)$ and $r \to \infty$ all the expectations in (3.21) tend to zero exponentially fast. In particular, we wish to drop the last expectation in (3.21) which is bounded above by

$$(3.22) \qquad E_r\{XI_{p_2}(r, X+1) I_{p_3}(r, X+1)\} \sim \frac{rq_1}{p_1} I_{p_2}(r, \frac{rq_1}{p_1}) I_{p_3}(r, \frac{rq_1}{p_1})$$

$$\sim C'r \frac{e^{-Cr}}{2\pi C''r} = \mathcal{O}(e^{-Cr})$$

where $0 < C = \frac{1}{2q_1} (1-\frac{p_2}{p_1})^2 + \frac{1}{2q_1} (1-\frac{p_3}{p_1})^2$; here the normal approximation and the first term of the Feller-Laplace expansion of the normal tail were used in the second line of (3.22). Using the normal approximation as in equation (4.8) of [6], we obtain for large r

$$(3.23) \qquad Z \sim \frac{rq_1}{p_1} + \frac{r(p_1-p_2)}{p_1p_2} \Phi\left(-\Delta_{21}\sqrt{\frac{r}{D_2}}\right) + \frac{r(p_1-p_3)}{p_1p_3} \Phi\left(-\Delta_{31}\sqrt{\frac{r}{D_3}}\right) ,$$

where $D_j = q_1p_j^2 + q_jp_1^2$ and $r = r(\Delta^*, P^*, k)$ was determined after (3.15) so as to satisfy (1.1). For small values of Δ^* we can take the first term alone in (3.23) as an estimate of Z. Hence for $r \to \infty$ (or $\Delta^* \to 0$) we have from (3.19)

$$(3.24)\qquad E\{N|R_I\} \sim \frac{rq_1}{p_1}\left(\sum_{i=1}^{k}\frac{1}{q_i}\right).$$

If we define the expected loss or risk $E\{L|R\}$ under procedure R by

$$(3.25)\qquad E\{L|R\} = \sum_{i=1}^{k}(p_1-p_i)\,E\{N_i|R\}$$

where $p_1 = \max_i p_i$, then from (3.18) we have for $r \to \infty$

$$(3.26)\qquad E\{L|R_I\} \sim \frac{rq_1}{p_1}\sum_{i=1}^{k}\left(\frac{p_1-p_i}{q_i}\right).$$

4. Procedure R_I' and Comparisons with Procedure R_I

Let R_I' denote the procedure that uses the same inverse-sampling termination rule as R_I together with the vector-at-a-time (VT) sampling rule. Ties are decided by randomization, i.e., we select one of the c contenders that reached r successes at the final stage using an independent experiment with probability $1/c$ for each.

To obtain the $P\{CS|R_I'\}$ we consider the event that on the mth stage (i.e., after m vectors of observations and not before) the best player A_1 has his r^{th} success $(r \le m)$ and each of the remaining A_i $(i \ge 2)$ has at most $r-1$ successes. Summing on m, we obtain

$$(4.1)\qquad P\{CS|R_I'\} - Q = \sum_{m=r}^{\infty}\binom{m-1}{r-1}p_1^r q_1^{m-r}\prod_{i=2}^{k}\left[\sum_{j=0}^{r-1}\binom{m}{j}p_i^j q_i^{m-j}\right]$$

$$= p_1^r\sum_{j=0}^{\infty}\binom{j+r-1}{j}q_1^j\prod_{i=2}^{k}\left[I_{q_i}(j+1,\ r)\right] = E_r\left\{\prod_{i=2}^{k} I_{q_i}(X+1,\ r)\right\},$$

where Q is the contribution to the $P\{CS|R_I'\}$ arising from randomization in the case of ties. Since each I_q-function in (4.1) is strictly increasing in q, we minimize the right side of (4.1) by setting $p_i = p_2$ $(i = 3, 4, \ldots, k)$; this does not prove that we have a minimum for the $P\{CS|R_I'\}$, although it is a proof for the asymptotic $(r \to \infty)$ case. To prove that $p_i = p_2$ $(i = 3, 4, \ldots, k)$ also yields a minimum of the $P\{CS|R_I'\}$ for small r, we write Q in the form

$$(4.2) \qquad \frac{1}{2}\{T_{1,2} + \cdots + T_{1,k}\} + \frac{1}{3}\{T_{1,2,3} + \cdots + T_{1,k-1,k}\} + \cdots + \frac{1}{k}\{T_{1,2,\ldots,k}\},$$

where, e.g., $T_{1,2}$ is the probability that A_1 and A_2 (and only these two) tie for first place by getting their r^{th} success on the same vector and before the others. Thus for the pair $(1, \alpha)$ with any $\alpha \neq 1$

$$(4.3) \qquad T_{1,\alpha} = p_1^r \sum_{j=0}^{\infty} \binom{j+r-1}{j} q_1^j \left[I_{q_\alpha}(j,r) - I_{q_\alpha}(j+1, r)\right] \prod_{\substack{i=2 \\ i\neq\alpha}}^{k} \left[I_{q_i}(j+1,r)\right],$$

for the triple $(1, \alpha, \beta)$ with $\alpha \neq \beta$ arbitrary (but not equal to 1)

$$(4.4) \qquad T_{1,\alpha,\beta} = p_1^r \sum_{j=0}^{\infty} \binom{j+r-1}{j} q_1^j \prod_{x=\alpha,\beta} \left[I_{q_x}(j,r) - I_{q_x}(j+1,r)\right] \prod_{\substack{i=2 \\ i\neq\alpha,\ i\neq\beta}}^{k} \left[I_{q_i}(j+1,r)\right],$$

etc. Multiplying the differences in square brackets and using (4.2) to combine terms, we find that a typical term has h factors of the form $I_{q_\alpha}(j+1, r)$ and k-1-h factors of the form $I_{q_\beta}(j,r)$, where α runs over a fixed set S_h of h values among $(2,3, \ldots, k)$ and β runs over the complementary set CS_h; let $\mathcal{S}_h$ denote the set of size $\binom{k-1}{h}$ consisting of all such sets S_h of size h. The coefficient W_h of this typical term, starting from the right end of (4.2), is

$$(4.5) \qquad W_h = (-1)^h\left\{\frac{1}{k} - \binom{h}{1}\frac{1}{k-1} + \binom{h}{2}\frac{1}{k-2} - \cdots + (-1)^h\binom{h}{h}\frac{1}{k-h}\right\}$$

$$= \int_0^1 \left(\frac{1}{x} - 1\right)^h x^{k-1} dx = \frac{1}{k\binom{k-1}{h}} > 0.$$

Hence we can write the exact value of the $P\{CS|R_I'\}$ in the form

$$(4.6) \qquad P\{CS|R_I'\} = p_1^r \sum_{j=0}^{\infty} \binom{j+r-1}{j} q_1^j \sum_{h=0}^{k-1} W_h \sum_{S_h \in \mathcal{S}_h} \left[\prod_{\alpha \in S_h} I_{q_\alpha}(j+1,\ r) \right] \left[\prod_{\beta \in CS_h} I_{q_\beta}(j,r) \right].$$

Since $W_h > 0$, all terms in (4.6) are positive and it follows as above that we minimize $P\{CS|R_I'\}$ by setting $p_i = p_2$ $(i = 3, 4, \ldots, k)$. This simplifies (4.6) considerably and a lower bound to the $P\{CS|R_I'\}$ for $p_i \geq p_2$ becomes

$$(4.7) \qquad \underset{\substack{p_i \geq p_2 \\ (i=3,4,\ldots,k)}}{\operatorname{Min}} P\{CS|R_I'\} = p_1^r \sum_{j=0}^{\infty} \binom{j+r-1}{j} q_1^j \frac{1}{k} \sum_{h=0}^{k-1} I_{q_2}^{h}(j+1,\ r)\, I_{q_2}^{k-1-h}(j,r)$$

$$= p_1^r \sum_{j=0}^{\infty} \binom{j+r-1}{j} q_1^j \frac{\left[I_{q_2}^k(j,r) - I_{q_2}^k(j+1,r) \right]}{k\left[I_{q_2}(j,r) - I_{q_2}(j+1,r) \right]} .$$

The same minimization can also be applied in (2.12) for procedure R_I and we clearly note that the result is exactly the same as in (4.7) above. Hence, after the first step of minimization, the $P\{CS\}$ expressions for R_I and R_I' are exactly the same. It follows that the least favorable configuration is the same for R_I and R_I' and that they require exactly the same value of r to satisfy (1.1).

To obtain the expected total number of observations $E\{N|R_I'\}$ under procedure R_I', we use the fact that we have an expression F_1 like (4.6) with the extra factor $km = k(j+r)$ if we select A_1 and $k-1$ similar expressions F_i corresponding to the selection of A_i $(i = 2, 3, \ldots, k)$. Thus

$$(4.8) \qquad F_1 = kp_1^r \sum_{j=0}^{\infty} (j+r)\binom{j+r-1}{j} q_1^j \sum_{h=0}^{k-1} W_h \sum_{S_h \in \mathcal{S}_h} \left[\prod_{\alpha \in S_h} I_{q_\alpha}(j+1,r) \right] \left[\prod_{\beta \in CS_h} I_{q_\beta}(j,r) \right]$$

$$= \frac{kr}{p_1} E_{r+1} \left\{ \sum_{h=0}^{k-1} W_h \sum_{S_h \in \mathcal{S}_h} \left[\prod_{\alpha \in S_h} I_{q_\alpha}(j+1,r) \right] \left[\prod_{\beta \in CS_h} I_{q_\beta}(j,r) \right] \right\}$$

where W_h is given by (4.5) the F_i $(i = 2, 3, \ldots, k)$ are obtained by interchanging p_1 with p_i (and q_1 with q_i), and

$$(4.9) \qquad E\{N|R_I'\} = \sum_{i=1}^{k} F_i,$$

To get an asymptotic approximation for (4.9) when $p_1 > p_i$ for $i \geq 2$, we first show that every F_i $(i \geq 2)$ tends to zero as $r \to \infty$. It suffices to show that for $q_2 > q_1$ and $r \to \infty$

$$(4.10) \qquad p_2^r \sum_{j=0}^{\infty} \binom{j+r-1}{j} q_2^j I_{q_1}(j,r) \sim I_{q_1}\left(\frac{rq_2}{p_2}, r\right) = o\left(\frac{1}{r}\right).$$

We interpret (4.10) as the probability that $Y_{p_1} \geq Y_{p_2}$ where $Y_{p_i} = X_{p_i}/r$ and X_{p_i} is the negative binomial with parameter p_i and common index r; the expectation of Y_{p_i} is q_i/p_i and the variance is $q_i/rp_i^2 \to 0$ $(i = 1, 2)$ as $r \to \infty$. Thus for an asymptotic $(r \to \infty)$ analysis we can replace Y_{p_2} by q_2/p_2 (or X_{p_2} by rq_2/p_2) and this gives the middle expression in (4.10). Using a normal approximation to the binomial as in (3.22), we obtain for $\Delta = p_1 - p_2 = q_2 - q_1 > 0$ and $r \to \infty$

$$(4.11) \qquad I_{q_1}\left(\frac{rq_2}{p_2}, r\right) \sim \Phi\left(\frac{-\Delta\sqrt{r}}{p_2\sqrt{q_2}}\right) \sim \frac{C_1}{r} e^{-C_2 r} = o\left(\frac{1}{r}\right) .$$

For the non-zero term F_1 in (4.9) we do a similar analysis and every I_q-function approaches 1 in expectation. Hence by (4.8) we obtain for

$$(4.12) \qquad E\{N|R_I'\} \sim \frac{kr}{p_1} .$$

To obtain the total expected number of observations from the non-best populations we replace k in (4.12) by $k-1$. Using the expected loss defined in (3.25) we obtain for $\Delta > 0$

$$(4.13) \qquad E\{L|R_I'\} = \frac{1}{k}\left[\sum_{i=1}^{k} F_i\right] \sum_{j=1}^{k} (p_1 - p_j) \sim \frac{r}{p_1} \sum_{j=1}^{k} (p_1 - p_j),$$

where the last expression holds for large r.

Since $q_1 < q_i$ $(i = 2, 3, \ldots, k)$, we find by comparing (3.24) and (4.12) that for large r the procedure R_I requires a uniformly smaller expected total number of trials when $\Delta > 0$. In addition, for large r procedure R_I has a uniformly smaller expected loss when $\Delta > 0$.

To approximate the value of r above which these results hold we now return to (3.17). A finer analysis of the application of (3.17) to (2.18) shows that a constant (with respect to r) is obtained from the omitted term in (3.17) whenever $\gamma = 1$ in (2.18). For any i, we find that $\gamma = 1$ in exactly $i - 1$ of the equations in (2.18), namely for $j = 2, 3, \ldots, i$ in the first line of (2.18). Moreover, for each i the contribution to $E\{N|R_I\}$ is $1/kq_i$. For $\gamma > 1$ we can use an argument similar to that in (4.11) to show that the omitted sums approach zero as $r \to \infty$. Hence we can replace (3.24) by the finer result

$$(4.14) \qquad E\{N|R_I\} = \frac{rq_1}{p_1}\left(\sum_{i=1}^{k} \frac{1}{q_i}\right) + \frac{1}{k}\sum_{i=1}^{k}\left(\frac{i-1}{q_i}\right) + o(1),$$

and a similar result holds for $E\{L|R_I\}$ if we replace q_i by $q_i/(p_1-p_i)$ for $i \geq 2$. For procedure R_I there are no corresponding non-zero terms omitted in (4.12) and (4.13). Hence we approximate the value of r above which the stated result for $E\{N\}$ holds by the solution in r of

$$(4.15) \qquad \frac{rq_1}{p_1}\left(\sum_{i=1}^{k} \frac{1}{q_i}\right) + \frac{1}{k}\sum_{i=1}^{k} \frac{i-1}{q_i} = \frac{kr}{p_1} .$$

For $p_2 = p_3 = \cdots = p_k$, this has the solution

$$(4.16) \qquad r = p_1/2\Delta ,$$

which is the same as that obtained in [2] for $k = 2$.

Similarly we approximate the value of r above which the stated result on $E\{L\}$ holds by the solution in r of

$$(4.17) \qquad \frac{rq_1}{p_1} \sum_{i=1}^{k} \left(\frac{p_1-p_i}{q_i}\right) + \frac{1}{k} \sum_{i=1}^{k} \frac{(i-1)(p_1-p_i)}{q_i} = \frac{r}{p_1} \sum_{i=1}^{k} (p_1-p_i).$$

For $p_2 = p_3 = \cdot \cdot \cdot = p_k$ this is the same equation as (4.15) and hence (4.16) again gives the required solution.

5. The Case of k Equal Success-Parameters.

Although a complete discussion of $E\{N\}$ requires consideration of the general case with exactly s success parameters equal in value to p_1 for $2 \le s \le k$ ($s = 1$ being already considered), we only consider the extreme case $s = k$, where all success parameters are equal.

For completeness we obtain some exact expressions from (2.18) and (4.9) but because of difficulties of analysis we shall not use these to form asymptotic expressions. From (2.18) for $q_1 = q_2 = \cdot \cdot \cdot = q_k = q$ (say)

$$(5.1) \qquad E\{N|R_I\} = \frac{1}{q} \sum_{\alpha=0}^{\infty} I_q(\alpha+1, r) \left[\frac{I_q^k(\alpha,r) - I_q^k(\alpha+1,r)}{I_q(\alpha,r) - I_q(\alpha+1,r)}\right]$$

which can be shown to converge for $0 \le q < 1$. For $q = 1$ the value is infinite as it should be since when all the p_i are zero (and only for this point) we never get r successes; hence with probability one the procedure R_I (as well as R_I') does not terminate for that one point. For $k = 1$ we note that (5.1) gives the correct result r/p and for $k=2$ it agrees with (2.15) of [6]. Using the right side of (2.9) it is easily seen that

$$(5.2) \qquad E\{N|R_I\} \sim \frac{k}{q} \sum_{\alpha=0}^{\infty} I_q^k(\alpha+1,r) \quad ,$$

but for $k \ge 2$ this does not directly lead to a simple asymptotic

$(r \to \infty)$ expression. Similarly, from (4.8) and (4.9) for a common q we use kF_1 to obtain

$$(5.3) \qquad E\{N|R_I'\} = \frac{kr}{p} E_{r+1}\left\{\frac{I_q^k(X, r) - I_q^k(X+1, r)}{I_q(X, r) - I_q(X+1, r)}\right\} \sim \frac{k^2 r}{p} E_{r+1}\left\{I_q^{k-1}(X, r)\right\},$$

but for $k \geq 3$ this does not give us a simple asymptotic $(r \to \infty)$ expression.

Starting with R_I', we use the fact that the expectation of the minimum of k independent negative-binomial (NB) chance variables, each with common success probability p and index r, is asymptotically $(r \to \infty)$ equivalent to the $100/(k+1)$ percentile of the underlying negative binomial distribution. Thus the expected total number of observations χ_p from each of the populations until any one of them reaches r successes is asymptotically $(r \to \infty)$ the solution in s of

$$(5.4) \qquad I_p(r, s) \equiv p^r \sum_{j=0}^{s-r} \binom{j+r-1}{j} q^j = \frac{1}{k+1} .$$

Since s and r will both be large, we use the normal approximation to the NB and replace (5.4) by

$$(5.5) \qquad P\left\{\frac{X_p - \frac{r}{p}}{\sqrt{rq}/p} < \frac{s - \frac{r}{p}}{\sqrt{rq}/p}\right\} \sim \Phi\left(\frac{sp-r}{\sqrt{rq}}\right) = \frac{1}{k+1} .$$

Multiplying the solution of (5.5) by k, we obtain a desired result

$$(5.6) \qquad E\{N|R_I'\} \sim ks = \frac{k}{p}(r - \lambda\sqrt{rq}) ,$$

where $\lambda = \lambda(k)$ is the $100k/((k+1)$ percentile of the standard normal distribution. For $k = 1$ we note that $\lambda = 0$ and (5.6) is exact.

For the procedure R_I we superimpose on the same data obtained a vector-at-a-time, the PWC procedure and note that the same population that terminated R_I' (by reaching r successes first) will also terminate R_I. This is because the number of failures from different populations can differ by at most one. Hence the asymptotic value of s for the winning population is the same and thus it also has sq failures. By PWC sampling, all the populations have sq failures and hence each of the $k-1$ non-winners have sp successes. Thus we again obtain

$$E\{N|R_I\} \sim ks = \frac{k}{p}(r - \lambda\sqrt{rq}) \tag{5.7}$$

and it follows that when p is common the procedures R_I and R_I' are asymptotically equivalent.

We have not proved for each r that the maximum of $E\{N\}$ for fixed p_1 occurs when all the p_i $(i \geq 2)$ are equal to p_1, but we note from (3.24) and (4.12) that this holds asymptotically $(r \to \infty)$ for both R_I and R_I'.

For the expected loss criterion with a common p and any r, we find that $E\{L\} = 0$ for both procedures. From (3.26) and (4.13) we note that the maximum for fixed p_1 may occur when the p_i $(i \geq 2)$ are equal, but not equal to p_1. In summary, the procedure R_I with PWC sampling is asymptotically $(r \to \infty)$ superior to R_I' with VT sampling throughout the parameter space with respect to both $E\{N\}$ and $E\{L\}$.

REFERENCES

[1] Bechhofer, R. E., Kiefer, J., Sobel, M. (1968). Sequential Identification and Ranking Problems, University of Chicago Press, Chicago.

[2] Gupta, S.S. (1963). Probability integrals of multivariate normal and multivariate t. Ann. Math. Statist. 34, 792-828.

[3] Milton, R.C. (1963). Tables of the equally correlated multivariate normal probability integral. Tech. Rep. No. 27, Department of Statistics, University of Minnesota, Minneapolis.

[4] Sobel, M., and Huyett, M., (1957). Selecting the best one of several binomial populations. Bell System Tech. J. 36, 537-576.

[5] Sobel, M., and Weiss, G. H. (1969). Play-the-winner sampling for selecting the better of two binomial populations. Univ. of Minnesota, Department of Statistics Technical Report No. 123.

[6] Sobel, M., and Weiss, G. H. (1969). Play-the-winner rule and inverse sampling in selecting the better of two binomial populations. Univ. of Minnesota Department of Statistics Technical Report No. 124.

APPENDIX

Another procedure R_I^* that is comparable with R_I and in some sense dual to it is defined by waiting until any one population has r failures (instead of r successes). Then that cycle is completed, so that under PWC sampling each population will have exactly r failures, and the population with the larger number of successes is declared to be the best. In case of ties we randomize between all contenders for first place. Since each population has exactly r failures at termination we can treat the populations separately and do not need the recursive-equation approach. The results are quite similar to those obtained above and it was therefore decided to include them here as an appendix.

Let Y_{p_i} $(i = 1, 2, \ldots, k)$ denote the random total number of observations required to obtain r failures from the population with success parameter p_i, where p_1 is the largest of the p_i and the rest are defined by the same cycle (starting with the best player A) as is used by the PWC sampling rule. Then for a population with arbitrary p

(A1) $$P\{Y_p = y\} = q^r \binom{y-1}{r-1} p^{y-r} \qquad y = r, r+1, \ldots ,$$

(A2) $$P\{Y_p < y\} = q^r \sum_{j=0}^{y-r-1} \binom{j+r-1}{j} p^j = I_q(r, y-r) \quad ,$$

the mean $E\{Y_p\} = r/q$ and the variance $\sigma^2(Y_p) = rp/q^2$. Hence the probability of a correct selection (CS) is given by

(A3) $$P\{CS|R_I^*\} - Q = q_1^r \sum_{j=0}^{\infty} \binom{j+r-1}{j} p_1^j \prod_{\alpha=2}^{k} I_{q_\alpha}(r, j) \quad ,$$

where Q is the contribution that arises from randomization when there are ties for first place. If we let T_{12} denote the probability that the two populations with parameters p_1 and p_2 (and only these two) tie for first place, etc., then

$$\text{(A4)} \qquad Q = \frac{1}{2}\{T_{1,2} + T_{1,3} + \cdots + T_{1,k}\} + \frac{1}{3}\{T_{1,2,3} + \cdots + T_{1,k-1,k}\} + \cdots + \frac{1}{k} T_{1,2,\ldots,k} ,$$

where, for example,

$$\text{(A5)} \qquad T_{1,2} = q_1^r \sum_{j=0}^{\infty} \binom{j+r-1}{j} p_1^j \, [I_{q_2}(r, j+1) - I_{q_2}(r, j)] \prod_{\alpha=3}^{k} I_{q_\alpha}(r, j)$$

and $T_{1,2,3}$ contains two such differences in square brackets, etc. We wish to show that all negative signs, as in (A5), disappear when we multiply out all the square brackets that arise. Consider any term that contains a fixed subset of h functions $I_{q_\alpha}(r, j+1)$ with argument $j+1$. For any $h (0 \le h \le k)$ and any subset of size h, the final coefficient which we denote by W_h will be

$$\text{(A6)} \qquad \sum_{i=0}^{k-h-1} \frac{(-1)^{k-h-i-1}}{k-i} \binom{k-h-1}{i} = \int_0^1 x^{k-1}\left(\frac{1}{x} - 1\right)^{k-h-1} dx = \frac{1}{k\binom{k-1}{h}} .$$

Hence, if we let S_h denote any fixed subset of size h, CS its complement, and $\mathcal{S}_h$ denote the $\binom{k-1}{h}$ possible subsets of size h, then we can write the exact $P\{CS|R_I^*\}$ as

$$\text{(A7)} \qquad P\{CS|R_I^*\} = q_1^r \sum_{j=0}^{\infty} \binom{j+r-1}{j} p_1^j \sum_{h=0}^{k-1} W_h \sum_{S_h \in \mathcal{S}_h} \left[\prod_{\alpha \in S_h} I_{q_\alpha}(r, j+1)\right] \left[\prod_{\beta \in CS_h} I_{q_\beta}(r, j)\right]$$

It immediately follows that in the minimization subject to $q_i \geq q_1 + \Delta^*$ we can set $q_i = q_1 + \Delta^* = \hat{q}$ (say) $(i = 2, 3, \ldots, k)$ and obtain

(A8) $$\text{Min } P\{CS|R_I^*\} = q_1^r \sum_{j=0}^{\infty} \binom{j+r-1}{j} p_1^j \frac{[I_{\hat{q}}^k(r, j+1) - I_{\hat{q}}^k(r, j)]}{k[I_{\hat{q}}(r, j+1) - I_{\hat{q}}(r, j)]} .$$

The exact expected total sample size is easily seen to be

(A9) $$E\{N|R_I^*\} = r \sum_{i=1}^{k} \frac{1}{q_i} .$$

To evaluate r so as to satisfy (1.1) with R_I replaced by R_I^*, we use the normal approximation as in (3.2) and obtain

(A10) $$\text{Min } P\{CS|R_I^*\} \sim \int_{-\infty}^{\infty} \Phi^{k-1}\left(\frac{x\sqrt{\rho} + H}{\sqrt{1-\rho}}\right) d\Phi(x) \equiv A_{k-1}(\rho, H)$$

where

(A11) $$H = \frac{\Delta^*\sqrt{r}}{\sqrt{p_1\hat{q}^2 + \hat{p}q_1^2}} \quad \text{and} \quad \rho = \frac{p_1\hat{q}^2}{p_1\hat{q}^2 + \hat{p}q_1^2} .$$

These results are similar to those in (3.4), but p and q are interchanged. The final minimization therefore leads to

(A12) $$q_1 = \frac{2}{3} - \frac{\Delta^*}{2} ; \quad q_2 = q_3 = \cdots = q_k = \frac{2}{3} + \frac{\Delta^*}{2} \; (= \hat{q}).$$

Putting this in the second expression in A(11) gives

(A13) $$\rho = \frac{1}{2} + \frac{3\Delta^*}{2} + \mathcal{O}\{(\Delta^*)^2\} ;$$

although $\rho = \frac{1}{2}$ now will provide a lower bound for small Δ^*, a correction term as in (3.14) is preferable. Using the lemma in (3.8) we now obtain

(A14) $$\text{Min } P\{CS|R_I^*\} \sim A_{k-1}\left(\tfrac{1}{2}, H\right) + \frac{3\Delta^*}{2}\frac{(k-1)(k-2)}{\sqrt{6}}\varphi(H)\varphi\left(\frac{H}{\sqrt{3}}\right)A_{k-3}\left(\frac{1}{4}, \frac{H}{\sqrt{6}}\right).$$

The first approximation for r is the solution of

(A15) $$A_{k-1}\left(\tfrac{1}{2}, H\right) = P^*$$

and if $\lambda = \lambda(P^*)$ is the table value of H that satisfies (A15), then

(A16) $$r \sim \frac{8}{27}\left(\frac{\lambda}{\Delta^*}\right)^2$$

is the first approximation for r and the equation setting (A14) equal to P^* can be then used to make minor corrections in r as in Section 3.

Hence $E\{N|R_I^*\}$ is given by (A9) with r replaced by the right side of (A16). Comparing with (3.24) we find that procedure R_I is preferred when

(A17) $$\frac{q_1}{p_1} < 1 \qquad \text{or} \qquad p_1 > \frac{1}{2}$$

and procedure R_I^* is preferred when $p_1 < \frac{1}{2}$.